PLAYING TEXTUAL ANALYSIS AS MUSIC*

George R S Weir
Department of Computer and Information Sciences,
University of Strathclyde, Glasgow G1 1XH, UK
george.weir@cis.strath.ac.uk

Marina Livitsanou
Department of Computer and Information Sciences,
University of Strathclyde, Glasgow G1 1XH, UK
mlivitsa@cis.strath.ac.uk

ABSTRACT

Textual analysis tools are now readily available and accessible to all but there are often difficulties in securing clear and useful analytical results, since the levels of abstraction employed in the tools is often mismatched to the level of interest of the researcher. Many tools provide low-level statistical analyses, with extensive data output, but leave the task of assimilating the significance of the results up to the investigator. In this setting, we have begun investigating a novel approach to data projection that may assist in a variety of assimilation contexts. We have developed a prototype software tool that takes the data output from an existing textual analysis program and maps the results to a musical 'interpretation'. This allows listeners to hear the textual characteristics of a particular document or to compare characteristics across several documents according to the musical projection of the data analyses.

1. INTRODUCTION

The use of software tools to analyse large texts or collections of texts is increasingly employed as a means of highlighting characteristics in the texts. High speed processing and the automated nature of such approaches mean that they often light on details that may be missed through manual analysis or casual observation. Highlighting differences or similarities between texts finds application in genre analysis, authorship studies and document forensics, in addition to other mainstream research domains such as historical linguistics, and literary studies. Although the accessibility of textual analysis tools is on the rise, there are often difficulties in securing clear analytical results, since the levels of abstraction employed in the tools is often mismatched to the level of interest of the researchers. For instance, many tools provide low-level statistical analyses, with extensive data output, but leave the task of assimilating the significance of the results up to the investigator. In this setting, we have begun investigating a novel approach to data projection that may assist in a variety of assimilation contexts.

What if we could compare documents by their 'sound'? Is it possible to categorize and distinguish documents, for example, by genre or by author, by the way they sound? Can we find a way to map characteristics of the text to characteristics of music such that the mapping will not only be meaningful (revealing) but will also produce sounds that are recognizably musical (and pleasant)?

To explore these questions, we have developed a prototype software tool that takes the data output from an existing textual analysis program and maps the results to a musical 'interpretation'. This allows listeners to hear the textual characteristics of a particular document or to compare the textual characteristics across several documents according to the musical projection of the data analyses.

Our primary purpose in this work is to explore the feasibility of establishing effective mappings between features in the textual analysis results and musical characteristics. As well as describing the approach that we have developed, we report on a small scale test to consider the effectiveness of using such musical rendering as a basis for genre recognition.

2. KEY OBJECTIVES

The main objective of our work was to design and implement an application that could take the textual analysis of sets of texts and convert them to audible form. In this process, the conversion must take account

* From *Corpus, ICT, and Language Education*, G R S Weir & S Ishikawa (Eds), Proceedings of the 6th International ICTATLL Conference, University of Strathclyde Press, 2010, pp. 111-118.

of the salient textual characteristics and the mapping of these characteristics to audible features must be performed in a way that minimises the loss of significant text information. Added to this key requirement on the textual features, is the need to produce results that are musical and pleasing to the ear. With these factors in mind, our main objective was to explore the effects of mapping between text characteristics and music characteristics and evaluate the feasibility of employing musical ‘interpretation’ of textual characteristics as a means of comparing and contrasting different texts.

The possibilities for mapping from textual characteristics to music are very extensive. There are many different variables in the textual domain and potentially as many in the musical domain. This provides us with an opportunity and a challenge. The opportunity lies in the range of different possibilities that may be explored as a basis for the musical projection of the textual data. The large mapping space in each domain affords innumerable different possible ways of ‘converting’ the textual features to musical features and this scope may increase the likelihood of one or more effective renderings across these two domains.

The challenge that arises from the large mapping space in each domain is to discover at least some of the mappings that satisfy our requirements: the results will be meaningful (revealing) but will also be recognizably musical (and pleasant).

For example, the number of words in a document can be mapped to the tempo of the music, or parts of speech can be mapped to different notes on the major scale, and so on. Whatever choices are taken, the selected mapping should accommodate the significant information available from the textual analysis and therefore, must be applied in a way that takes advantage of the range of values for these characteristics.

The mapping challenge may be highlighted through a further example. Suppose that the average sentence length for a set of analysed texts ranges between 4 and 24. That means 20 different values can be derived directly for this characteristic. We may consider mapping this feature of the texts to the key signature for their musical interpretation. This would have the aim of discriminating between the text samples – according to their average sentence length – by the resultant key for each of their musical renderings.

This mapping is problematic since the key signature for a piece of music may take seventeen different values. Therefore, in this case, if we try to map the average sentence length of the documents to the key signatures of their music, this is not ideal since we are trying to fit 20 values to 17. A consequence of adopting this component in our mapping of these example texts is that some information about our texts is lost. Ideally, we need to find and map characteristics that have the same value range.

Of course, the mapping can consider many different ‘dimensions’ from the textual data, of which average sentence length is just one, and a desirable mapping should accommodate all key variable ranges in the texts. Another possibility would be to map according the levels of information. For example, a more general text characteristic (e.g., number of words) can be mapped to a general music characteristic (e.g., tempo). The value range constraint would apply to this case as well, since we do not want to lose document information.

Vying with all of these considerations is our further constraint that the result from the selected set of mappings should be acceptable as music (i.e., with all the main musical characteristics: tempo, key signature, and note variations within the same scale) and should sound pleasant. This is perhaps the most challenging aspect of this work.

3. RELATED WORK

3.1 Musical Analysis

Musical analysis can help to understand the structure of a musical piece and expose how the music achieves its effects. As with textual analysis, musical analysis may also help to distinguish between musicians or among different genres of music by determining the presence of specific musical characteristics.

Music analysis techniques include ‘discretization’, which analyses a music piece by breaking it down into smaller and simpler parts and then examines the way that these parts fit together and interact with each other, and ‘compositional analysis’, in which music is analysed from a compositional viewpoint. Another area where musical analysis is applicable is in Music Information Retrieval (MIR). This deals with musical analysis as a means to retrieve information from music. Most applications of MIR are content-based, i.e., involve constructing a set of features for a text or an image or a music piece by analysing its actual contents (Orio, 2006).

Although our purpose in the present work is not to perform musical analyses, the available scope for such decomposition, affords a setoff 'primitive' components that can be used as the musical side of the mapping. We will return to these musical features shortly. In the meantime, we turn to a brief overview of textual analysis.

3.2 Textual Analysis

Textual analysis can be used in corpus linguistics to help with a variety of language studies. Most commonly, textual analysis is applied in order to analyse and compare documents. Such analysis may offer insights on the writing styles of different authors or the intrinsic characteristics of document genres.

Among the freely available software tools that can be used for textual analysis are Part Of Speech (POS) taggers, such as CRFTagger (Phan, 2006). POS taggers analyse text files by categorising the words according to their part of speech (e.g. nouns, verbs, adjectives and so on), but part of speech is only one of the textual features that analysis tools may annotate, quantify and record. Broad approaches, covering a wider variety of textual features, include tools such as Wordsmith (Scott, 1998), AntConc (Anthony, 2005) and the Posit Text Profiling Toolkit (Weir, 2007). More ambitious facilities are freely available in systems such as Nooj (Silberztein, 2005), and GATE (Bontcheva et al., 2004), which are both regularly updated. In addition, NLTK (Bird, 2006) provides a set of extensible programming modules primarily aimed at teaching natural language processing. (A useful review of several textual analysis programmes is provided by Alexa and Zuell, 1999).

A notable feature of several of textual analysis tools is that they seek to provide approachable means whereby non-computer specialists may analyse their own collections of textual data. This, in turn, creates a scenario in which the users may find difficulty in 'making sense' of the resultant analysis data. One case in point is presented in Table 1, below.

Table 1. Example summary data from Posit Toolset (Weir, op. cit.)

Input filename	emma.txt
Total words (tokens)	159826
Total unique words (types)	7364
Type/Token Ratio (TTR)	21.7037
Number of sentences	8585
Average sentence length (ASL)	18.6169
Number of characters	914519
Average word length (AWL)	5.72197
NUMBER OF TOKEN TYPES	
noun_types	4268
verb_types	2603
adjective_types	1346
adverb_types	487
preposition_types	65
personal_pronoun_types	23
determiner_types	18
possessive_pronoun_types	7
interjection_types	5
particle_types	0
NUMBER OF POS TYPES	
nouns	69060
verbs	67678
prepositions	38600
personal pronouns	31192
determiners	26178
adverbs	25432
adjectives	25086
possessive pronouns	9582
interjections	516
particles	0

This example shows ‘summary’ data for the text version of Jane Austin’s novel ‘Emma’, including the totals for tokens, types, part-of-speech types and tokens, average sentence length and type/token ratio. Such data is easily generated from such software tools but interpretation is less straightforward. For instance, using such analysis tools, we may readily generate and review data for all of the novels by Jane Austin and this may reveal some commonalities across the texts. Yet, such comparisons, whether from the data tables or from graphed data projections, are not straightforward. As with many contexts of data analysis, the quantity of data generated through the analysis makes it difficult to see the wood for the trees.

4. APPROACHES TO MAPPING

As indicated above, there are many possible factors that may be mapped from textual data to musical feature. To date, we have only explored a small sub-set of these possibilities, as a means to determine the feasibility of musical interpretation of textual analysis data. In the following, we present examples of the selected mappings that we have adopted in the initial stages of this process.

Below there is a description of the mapping of each textual characteristic to musical characteristic along with an explanation for the motivation behind it.

In our first approach, we have the average sentence length mapped to the tempo. This was based on the likely range of the values that these characteristics can take. In order to evaluate approximately the possible values for the average sentence length we analysed a variety of documents and noted the average sentence length values from these documents. Taking the highest and the lowest values, we calculated the range which was 18. This range maps exactly to the range available for musical tempo. Although, in principle, the tempo of a musical piece could take any value, in practice this is not the case. There is no piece with a tempo of ten beats per minute, since this would be much too slow, nor any musical piece with a tempo of 500 beats per minute, since this would be too fast. Hence, through similar musical analysis, we narrowed the range of the values for the tempo to 18.

The type/token ratio (unique words / total words) was mapped to the key signature and this was once more based on the range of the values that these characteristics may have. A key signature can have 30 possible combinations (15 for major scale and 15 for minor scale). This was found to match the range of values for our sample of type token ratios. Once the key signature was determined, the next step was to find a mapping for the notes (pitch). This musical factor has a significant effect on the listener’s perception of the music and for this reason we considered two alternative mappings.

4.1 Melody Mapping 1: Parts of Speech

This mapping attends to the part of speech of each word and maps the notes to the words on this basis. At its simplest, nouns can be mapped to the C note (tonic); verbs can be mapped to the D note (supertonic) and so on. In our adopted approach to melody mapping via parts of speech, we took the ratios of the different parts of speech and used this to determine the note value. We chose the ratios and not the absolute values for parts of speech because the ratio affords a clearer distinction between documents. This can be understood better through the following simplified example:

Suppose we wish to compare a poem and a novel. The poem has 50 nouns and 20 verbs and a total of 200 words. Coincidentally, the novel has 50 nouns and 20 verbs and 500 total words. If the melody mapping was based on the absolute values of the nouns and verbs, the two documents above would have the same melody even though the first is a poem and the second is fiction. Working from the ratios, the musical result is different since the ratios of nouns and verbs are different.

In order to ensure that the musical result was sufficiently ‘interesting’, we added variable durations to notes and rests. Note duration was based on variations of the parts of speech (e.g., proper nouns get quarter duration, plural nouns get half duration and so on). Musical rests and their duration were based upon the punctuation marks within the analysed texts. Full stops were mapped to a one beat rest; colons to a half beat rest; commas to a quarter beat rest, and so on. The idea behind this was that the flow of the music should follow the flow of the document (i.e. when we read something we stop at full stops longer than we stop at commas, hence the music should pause longer for full stops than for commas).

One further feature of this melody mapping approach is that each word maps to a single note. As a result, the length of the document is associated directly with the length of the musical piece.

4.1 Melody Mapping 2: Word Frequency Ratio

In order to explore alternative mappings, melody mapping 2 derives the melody notes from the ratio of the frequencies of the words (i.e. the value of the ratio is the value of the note). Again, the mapping is based on the ratios and not the absolute values for the reasons explained above.

Melody mapping 2 also adds rests to the music according to the punctuation marks. For some punctuation marks though, such as question mark and exclamation mark instead of putting a rest, put a note and changed its attack velocity (how 'hard' the note will sound). The was motivated by the way that we change our voice when we read; for question marks we alter the tone of our voice, for exclamation marks we alter it even more. Accordingly, the notes for the exclamation marks have a bigger attack velocity than the notes for the question marks.

5. PLAYTEXT

A key component in the work reported here was the creation of a software application to convert the textual analysis data from a document to music. Our prototype, called PlayText, can create midi files directly from textual analysis data where the required data formats input to PlayText have been pre-processed using the Posit tools.

When the PlayText program starts (Figure 1), the play and stop buttons are disabled to indicate that a document must first be selected. When the user selects the title of a document from the appropriate pull down menu (representing a predefined set of example data), the analysis data files for the specified document are opened and read. At the same time, the textual content of the selected document is displayed in the user interface.

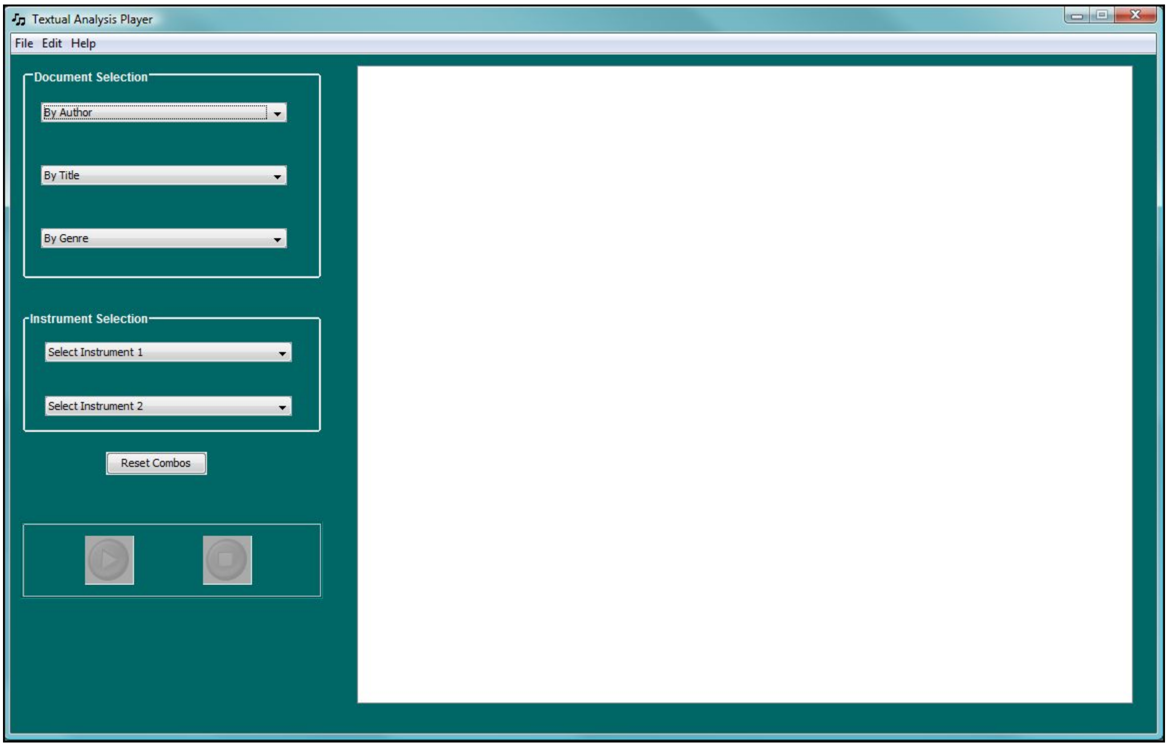

Fig. 1. User interface for Textual Analysis Player (PlayText)

Selecting the title, results in the authors and genres fields being updated accordingly. When the user clicks the play button, it changes to a pause button and the mapping is started. Depending on how long it will take to complete the mapping and generation of the midi file, a progress bar is displayed indicating the work is being done in the background. The resultant music starts playing as soon as the mapping is complete.

Figure 2 illustrates the program state after loading the data analysis for Jules Verne's 'Twenty Thousand Leagues Under the Sea'. This demonstrates the updated author and genre fields and the enabled Play button.

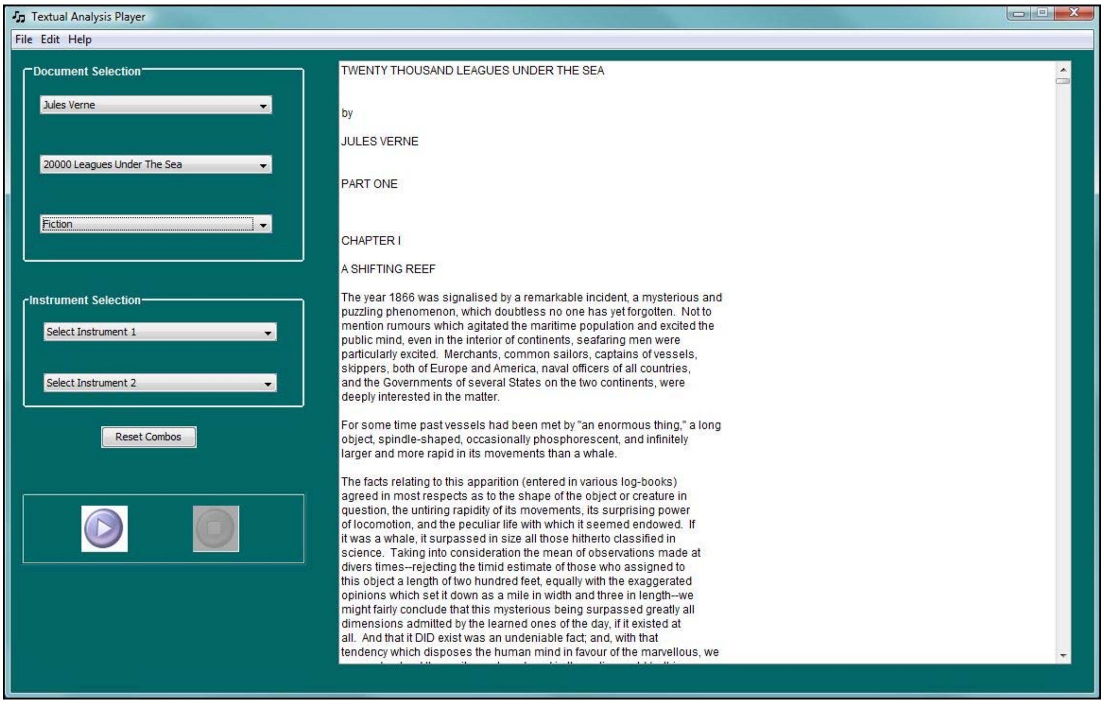

Fig. 2. PlayText with data loaded

The program offers a variety of additional features such as: open an existing midi, save a midi file and import a document. These features can be accessed through the File menu. The Edit menu affords features such as: change the font of the displayed document, add/remove the optional chords of a music piece, add/remove percussion, change the tempo of a music piece and choose between the two available mappings (by default the first mapping is selected). The user can also select the instruments for the main melody and the chords (by default these are both set to piano). Finally, the 'Reset Combos' button restores all of the default options so that the user can restart their selection.

6. EXPERIMENT

Using textual analysis data generated from text documents by the Posit tools, we applied our prototype Playtext system to produce sets of musical interpretations in each of the two melody mappings. With these examples, we then designed a small experiment to gauge whether our melody mappings would have an impact upon users' ability to discriminate textual data on the basis of its musical interpretation. In this experiment, the testers (8 subjects) were given six midi files from each mapping, and a form to complete for each mapping. Each of the midi files in a set was generated from different documents and the experiment had two components. In the first component, the testers, listened to the midi files and tried to 'guess' the genre of the document they came from. For example, midi file 1 is fiction; midi file 2 is poetry and so on. Six possible genres were listed in the form provided to the subjects: fiction, poetry, mythology, drama, history, and horror tales.

Although tenuous, we were interested to determine whether the music produced by the program suggested a document genre to the listener. In other words, can the mapping convey such information about the document?

The second exercise was to listen again to the same midi files only this time try to 'guess' which files, if any, were from the same author. Subjects were not given any suggested authors nor were they required to propose an author for any document. The aim was to see whether documents from the same author had any discernable musical similarity. The documents for the 1st mapping were:

1) John Keats - Endymion - Poetry
2) Edgar Alan Poe - The Fall of the House of Usher - Horror Tales
3) Charles Dickens - A Christmas Carol - Fiction
4) Bram Stoker - Dracula - Horror Tales
5) Edgar Alan Poe - The Raven - Poetry
6) Oscar Wilde - The Picture of Dorian Gray - Fiction

The documents for the 2nd mapping were:

1) William Shakespeare - Hamlet - Drama
2) Jules Verne - Around the World in 80 Days - Adventure
3) Homer - Iliad - Mythology
4) William Shakespeare - Othello - Drama
5) Edgar Alan Poe - A Dream within a Dream - Poetry
6) Gaston Leroux - The Phantom of the Opera - Fiction

5.1 Experimental Results

In the first part of the experiment (genres), mapping 1 allowed the subjects to correctly classify 4 out of the 6 midi files. Mapping 2 allowed the subjects to correctly classify 3 out of 6 midi files. In the second part of the experiment (authors), using mapping1, one subject correctly identified the midi files that came from documents of the same author. For mapping2, no one found the midi files that came from documents of the same author. Overall, mapping1 seemed to be more 'successful' in distinguishing the documents.

7. CONCLUSION

In this paper we have described our development of a prototype system (PlayText) that takes textual analysis data (generated using the Posit Textual Analysis Toolset) and maps textual characteristics to musical features. With this approach, we have produced two text-to-music mappings and undertaken a small experiment to determine whether such musical interpretations of texts can afford a novel means of discriminating between texts. Although only two of all possible mappings were employed experimentally, the results suggest that our approach goes some way toward achieving our original two objectives, viz. (1) map from text to music in a manner that aids discrimination and (2) produce musical results that are plausible as music. On the basis of our work so far, we conclude that there is considerable potential for further development, especially in terms of alternative text to music mapping. Furthermore, we propose that such an approach does offer a viable and novel means of projecting textual characteristics to musical characteristics in a manner that aids discrimination by the listener.

Several examples of the musical interpretations produced in the course of these trials are available for download as midi files from the following URL: http://personal.cis.strath.ac.uk/~gw/PlayText/index.html.

REFERENCES

Alexa M. & Zuell C. (1999). A Review of Software for Text Analysis. Mannheim, Zuma.

Anthony, L. (2005). AntConc: A Learner and Classroom Friendly, Multi-Platform Corpus Analysis Toolkit. *Proceedings of IWLeL 2004: An Interactive Workshop on Language e-Learning*.

Bird, S. (2006). NLTK: the natural language toolkit. In *Proceedings of the COLING/ACL on interactive Presentation Sessions.* Association for Computational Linguistics, Morristown, NJ, 69-72.

Bontcheva, K., Tablan, V., Maynard, D. & Cunningham, H. (2004). Evolving GATE to meet new challenges in language engineering. *Natural Language Engineering*, 10: 349-373.

Phan, X-H. (2006). CRFTagger: CRF English POS Tagger, http://crftagger.sourceforge.net/.

Orio, N. (2006). Music Retrieval: A Tutorial and Review, *Foundations and Trends in Information Retrieval*: Vol. 1: No 1, pp 1-90.

Scott, M. (1998). *Wordsmith Tools Version 3*, Oxford University Press, Oxford, UK.

Silberztein, M. (2005). NooJ: A Linguistic Annotation System for Corpus Processing. In *Proceedings of HLT/EMNLP on Interactive Demonstrations*, Vancouver, 10-11.

Weir, G. (2007). The Posit Text Profiling Toolset. *Proceedings of the 12th Conference of Pan-Pacific Association of Applied Linguistics.* Pan-Pacific Association of Applied Linguistics, Tokyo, Japan.